

Project Data Analysis for Retail: Sales Performance Report

Dataset Brief

Dataset yang digunakan berisi transaksi dari tahun **2009 sampai dengan tahun 2012** dengan jumlah raw data sebanyak **5500**, termasuk di dalamnya order status yang terbagi menjadi **order finished, order returned** dan **order cancelled**

Adapun dataset yang sudah diberikan dan akan digunakan pada project ini berisi data sebagai berikut.

1. OrderID
2. Order Status
3. Customer
4. Order Date
5. Order Quantity
6. Sales
7. Discount %
8. Discount
9. Product Category
10. Product Sub-Category

order_id	order_status	customer	order_date	order_quantity	sales	discount	discount_value	product_category	product_sub_category
----------	--------------	----------	------------	----------------	-------	----------	----------------	------------------	----------------------

Nama tabel yang akan digunakan pada project ini adalah **dqlab_sales_store**

Petunjuk Penyelesaian Project

Dari data yang sudah diberikan, dari pihak manajemen DQLab store ingin mengetahui:

1A. Overall performance DQLab Store dari tahun 2009 - 2012 untuk jumlah order dan total sales order finished.

1B. Overall performance DQLab by subcategory product yang akan dibandingkan antara tahun 2011 dan tahun 2012.

2A. Efektifitas dan efisiensi promosi yang dilakukan selama ini, dengan menghitung burn rate dari promosi yang dilakukan overall berdasarkan tahun.

2B. Efektifitas dan efisiensi promosi yang dilakukan selama ini, dengan menghitung burn rate dari promosi yang dilakukan overall berdasarkan sub-category.

Setelah melihat hasil analisa di Sub Bab 1 dan 2, selanjutnya dilakukan analisa terhadap customer DQLab.

3A. Analisa terhadap customer setiap tahunnya

1.DQLab Store Overall Performance

Pekerjaan pertama yang harus dilakukan adalah melakukan analisis terhadap overall performance dari DQLab Store berdasarkan tahun dan product sub category.

1.A Overall Performance by Year

Membuat Query dengan menggunakan SQL untuk mendapatkan total penjualan (sales) dan jumlah order (number_of_order) dari tahun 2009 sampai 2012 (years).

SQL Query :

```
1 select year(order_date) as years,sum(sales) as sales,count(*) as number_of_order
   from dqlab_sales_store where order_status='Order Finished' and year(order_date)
   between 2009 and 2012 group by year(order_date);
```

Output yang harus dihasilkan :

years	sales	number_of_order
2009	4613872681	1244
2010	4059100607	1248
2011	4112036186	1178
2012	4482983158	1254

1.B Overall Performance by Product Sub Category

Membuat Query dengan menggunakan SQL untuk mendapatkan total penjualan (sales) **berdasarkan sub category dari produk** (product_sub_category) pada **tahun 2011 dan 2012** saja (years) .

SQL Query :

```
1 SELECT
2     year (order_date) as years,
3     product_sub_category,
4     sum(sales) as sales
5 FROM
6     dqlab_sales_store
7 WHERE
8     year (order_date) between 2011 and 2012 and
9     order_status = 'Order Finished'
10 GROUP BY
11     year (order_date),
12     product_sub_category
13 order by years,sales desc
```

Output yang harus dihasilkan :

years	product_sub_category	sales
2011	Chairs & Chairmats	622962720
2011	Office Machines	545856280
2011	Tables	505875008
2011	Copiers and Fax	404074080
2011	Telephones and Communication	392194658
2011	Binders and Binder Accessories	298023200
2011	Storage & Organization	285991820
2011	Appliances	272630020
2011	Computer Peripherals	232677960
2011	Bookcases	169304620
2011	Office Furnishings	160471500
2011	Paper	111080380
2011	Pens & Art Supplies	43093800
2011	Envelopes	36463900
2011	Labels	15607780

2.DQLab Store Promotion Effectiveness and Efficiency

Setelah melihat overall performance dari DQLab Store.pekerjaan selanjutnya yang harus dilakukan adalah melakukan analisis keterkaitan antara sales dengan promosi yang sudah dilakukan oleh DQLab Store sebelumnya.

2.A Promotion Effectiveness and Efficiency by Years

Pada bagian ini kita akan melakukan analisa terhadap efektifitas dan efisiensi dari promosi yang sudah dilakukan selama ini

Efektifitas dan efisiensi dari promosi yang dilakukan akan dianalisa berdasarkan Burn Rate yaitu dengan membandingkan total value promosi yang dikeluarkan terhadap total sales yang diperoleh

DQLab berharap bahwa burn rate tetap berada diangka maksimum 4.5%

Formula untuk burn rate : $(\text{total discount} / \text{total sales}) * 100$

Buatkan Derived Tables untuk menghitung total sales (sales) dan total discount (promotion_value) berdasarkan tahun(years) dan formulasikan persentase burn rate nya (burn_rate_percentage).

SQL Query :

```
1 select year(order_date) years,sum(sales) sales,sum(discount_value)
   promotion_value,round(sum(discount_value)/sum(sales)*100,2) burn_rate_percentage from
   dqlab_sales_store where order_status='Order Finished' group by year(order_date);
```

Output yang harus dihasilkan :

years	sales	promotion_value	burn_rate_percentage
2009	4613872681	214330327	4.65
2010	4059100607	197506939	4.87
2011	4112036186	214611556	5.22
2012	4482983158	225867642	5.04

2.B Promotion Effectiveness and Efficiency by Product Sub Category

Pada bagian ini kita akan melakukan analisa terhadap efektifitas dan efisiensi dari promosi yang sudah dilakukan selama ini seperti pada bagian sebelumnya.

Akan tetapi, ada kolom yang harus ditambahkan, yaitu : product_sub_category dan product_category.

SQL Query :

```
1 select year(order_date) years,product_sub_category,product_category,sum(sales) as sales,sum(discount_value)
   as promotion_value,round(sum(discount_value)/sum(sales)*100,2)as burn_rate_percentage from dqlab_sales_store
   where order_status='Order Finished'and year(order_date)=2012 group by
   year(order_date),product_sub_category,product_category
2 order by sales desc ;
```

Output yang harus dihasilkan :

years	product_sub_category	product_category	sales	promotion_value	burn_rate_percentage
2012	Office Machines	Technology	811427140	46616695	5.75
2012	Chairs & Chairmats	Furniture	654168740	26623882	4.07
2012	Telephones and Communication	Technology	422287514	18800188	4.45
2012	Tables	Furniture	388993784	16348689	4.20
2012	Binders and Binder Accessories	Office Supplies	363879200	22338980	6.14
2012	Storage & Organization	Office Supplies	356714140	18802166	5.27
2012	Computer Peripherals	Technology	308014340	15333293	4.98
2012	Copiers and Fax	Technology	292489800	14530870	4.97
2012	Appliances	Office Supplies	266131100	14393300	5.41
2012	Office Furnishings	Furniture	178927480	8233849	4.60
2012	Bookcases	Furniture	159984680	10024365	6.27
2012	Paper	Office Supplies	126896160	6224694	4.91
2012	Envelopes	Office Supplies	58629280	2334321	3.98
2012	Pens & Art Supplies	Office Supplies	43818480	2343501	5.35
2012	Scissors, Rulers and Trimmers	Office Supplies	36776400	2349280	6.39
2012	Labels	Office Supplies	10007040	452245	4.52
2012	Rubber Bands	Office Supplies	3837880	117324	3.06

3.Customer Analytics

Setelah melihat overall performance yang tidak mengalami pertumbuhan yang signifikan, manajemen ingin mengetahui pertumbuhan dari sisi customer.

3.A Customers Transactions per Year

DQLab Store ingin mengetahui jumlah customer (number_of_customer) yang bertransaksi setiap tahun dari 2009 sampai 2012 (years).

SQL Query :

```
1 select year(order_date) years,count(distinct customer) number_of_customer
   from dqlab_sales_store where order_status='Order Finished'
2 group by year(order_date);
```

Output yang harus dihasilkan :

years	number_of_customer
2009	585
2010	593
2011	581
2012	594

Data Mentor

Nelda Ampulembang Parenta

Senior Data AnalystLogisly